A Fuzzy System for Concept-Level Sentiment Analysis

Mauro Dragoni, Andrea G. B. Tettamanzi, Célia da Costa Pereira

A Fuzzy System For Concept-Level Sentiment Analysis

Mauro Dragoni[1], Andrea G.B. Tettamanzi[2], and Célia da Costa Pereira[2]

[1] FBK–IRST, Trento, Italy
[2] Université Nice Sophia Antipolis, I3S, UMR 7271, Sophia Antipolis, France
dragoni@fbk.eu andrea.tettamanzi|celia.pereira@unice.fr

Abstract. An emerging field within Sentiment Analysis concerns the investigation about how sentiment concepts have to be adapted with respect to the different domains in which they are used. In the context of the Concept-Level Sentiment Analysis Challenge, we presented a system whose aims are twofold: (i) the implementation of a learning approach able to model fuzzy functions used for building the relationships graph representing the appropriateness between sentiment concepts and different domains (Task 1); and (ii) the development of a semantic resource based on the connection between an extended version of WordNet, SenticNet, and ConceptNet, that has been used both for extracting concepts (Task 2) and for classifying sentences within specific domains (Task 3).

1 Introduction And Related Work

Sentiment Analysis is a kind of text categorization task that aims to classify documents according to their opinion (polarity) on a given subject [1]. This task has created a considerable interest due to its wide applications. However, in the classic Sentiment Analysis the polarity of each term of the document is computed independently of the domain which the document belongs to. Recently, the idea of adapting term polarities to different domains emerged [2]. The rationale behind the idea of such investigation is simple. Let's consider the following example concerning the adjective "small":

1. The sideboard is **small** and it is not able to contain a lot of stuff.
2. The **small** dimensions of this decoder allow to move it easily.

In the first text, we considered the Furnishings domain and, within it, the polarity of the adjective "small" is, for sure, "negative" because it highlights an issue of the described item. On the other hand, in the second text, where we considered the Electronics domain, the polarity of such adjective can be considered "positive".

In the literature, different approaches related to the Multi-Domain Sentiment Analysis have been proposed. Briefly, two main categories may be identified: (i) the transfer of learned classifiers across different domains [3] [4], and (ii) the use of propagation of labels through graph structures [5] [6]. Independently of the kind of approach, works using concepts rather than terms for representing different sentiments have been proposed.

Unlike the approaches already discussed in the literature, we address the multi-domain sentiment analysis problem by applying the fuzzy logic theory for modeling membership functions representing the relationships between concepts and domains.

Moreover, the proposed system exploits the use of semantic background knowledge for propagating information represented by the learned fuzzy membership functions to each element of the network. To the best of our knowledge, the proposed approach is innovative with respect to the state of the art of Multi-Domain Sentiment Analysis.

The paper is structured as follows. Section 2 introduces the background knowledge and tools used during the development of the system that is described in detail in Section 3. Finally, Section 4 provides a description of how the tasks of the challenge have been addressed and concludes the paper.

2 Preliminaries

The system is implemented on top of background knowledge used for representing the linguistic connections between "concepts" described in several resources. Below, it is possible to find the list of such resources and the links where further information about them may be found.

WordNet[3] [7] is one of the most important resources available to researchers in the field of text analysis, computational linguistics, and many related areas. In the implemented system, WordNet has been used as a starting point for the construction of the semantic graph used by the system (see Section 3) However, due to some coverage limitations occurring in WordNet, it has been extended by linking further terms coming from the Roget's Thesaurus [8].

SenticNet[4] [9] is a publicly available resource for opinion mining, which exploits both Artificial Intelligence and Semantic Web techniques to infer the polarity associated with common-sense concepts and represent it in a semantic-aware format. In particular, SenticNet uses dimensionality reduction to calculate the affective valence of a set of Open Mind concepts and represent it in a machine-accessible and machine-processable format.

All resources have been connected by exploiting links contained in ConceptNet[5] [10] in order to build a single graph for representing the entire background knowledge exploitable by the system.

3 System

The main aim of the implemented system is the learning of fuzzy membership functions representing the degree of membership of a concept to a domain in terms of both sentiment polarity and aboutness. The two pillars on which the system has been built are: (i) the use of fuzzy logic for modeling the polarity of a concept with respect to a domain as well as its aboutness, and (ii) the creation of a two-levels graph where the top level represents the semantic relationships between concepts and the bottom level contains the links between all concept membership functions and the domains.

[3] https://wordnet.princeton.edu/
[4] http://sentic.net/
[5] http://conceptnet5.media.mit.edu/

Figure 1 shows the conceptualization of the two-levels graph. Relationships between the concepts of the Level 1 (the Semantic Level) are described by the background knowledge exploited by the system as described in Section 2. The type of relationships are the same generally used in linguistic resource: for example, concepts C_1 and C_3 may be connected through an Is-A relationship rather than the Antonym one. Instead, each connection of the Level 2 (the Sentiment Level) describes the membership of each concept in the different domains considered.

The system has been trained by using the Blitzer dataset[6] in two steps: first, the fuzzy membership functions have been initially estimated by analyzing only explicit information present within the dataset (Section 3.1); then, (ii) explicit information has been propagated through the Sentiment Level graph by exploiting the connections defined in the Semantic Level.

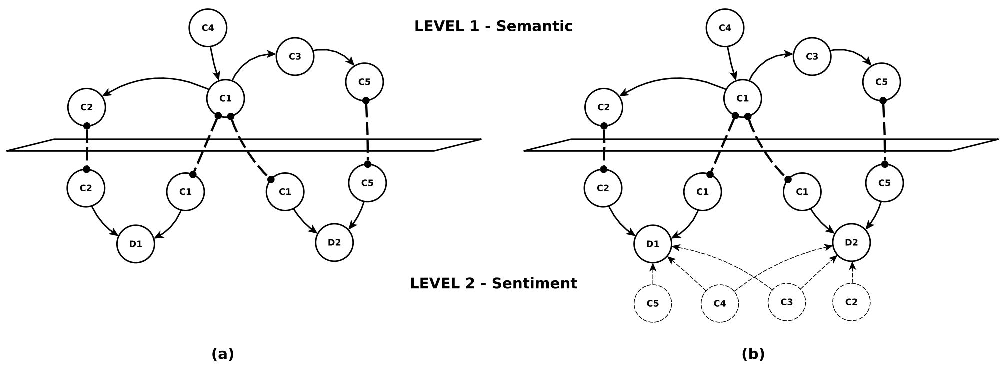

Fig. 1: The two-layer graph initialized during the Preliminary Learning Phase (a) and its evolution after the execution of the Information Propagation Phase (b).

3.1 Preliminary Learning Phase

The Preliminary Learning (PL) phase aims at estimating the initial polarity of each concept with respect to a domain. The estimation of this value is done by analyzing only explicit information provided by the training set. This phase allows to define the preliminary fuzzy membership functions between the concepts defined in the Semantic Level of the graph and the domains that are defined in the Sentiment Level. Such a value is computed by Equation 1:

$$\text{polarity}^*_i(C) = \frac{k^i_C}{T^i_C} \in [-1, 1] \qquad \forall i = 1, \dots, n, \tag{1}$$

where C is the concept taken into account, index i refers to domain D_i which the concept belongs to, n is the number of domains available in the training set, k^i_C is the

[6] http://www.cs.jhu.edu/ mdredze/datasets/sentiment/

arithmetic sum of the polarities observed for concept C in the training set restricted to domain D_i, and T_C^i is the number of instances of the training set, restricted to domain D_i, in which concept C occurs. The shape of the fuzzy membership function generated during this phase is a triangle with the top vertex in the coordinates $(x, 1)$, where $x = \text{polarity}_i^*(C)$, and with the two bottom vertices in the coordinates $(-1, 0)$ and $(1, 0)$, respectively. The rationale is that while we have one point (x) in which we have full confidence, our uncertainty covers the entire space because we do not have any information concerning the remaining polarity values.

3.2 Information Propagation Phase

The Information Propagation (IP) phase aims at exploiting the explicit information learned in the PL phase in order to both (i) refine the fuzzy membership function of the known concepts, as well as (ii) to model such functions for concepts that are not specified in the training set, but that are semantically related to the specified ones. Figure 1 presents how the two-levels graph evolves before and after the execution of the IP phase. After the PL phase only four membership functions are modeled: C_1 and C_2 for the domain D_1, and C_1 and C_5 for the domain D_2 (Figure 1a). However, as we may observe, in the Semantic Level there are concepts that are semantically related to the ones that were explicitly defined in the training set, namely C_3 and C_4; furthermore, there are also concepts for which a fuzzy membership function has not been modeled for some domains (i.e. C_2 for the domain D_2 and C_5 for the domain D_1).

Such fuzzy membership functions may be inferred by propagating the information modeled in the PL phase. Similarly, existing fuzzy membership functions are refined by the influence of the other ones. Let's consider the polarity between the concept C_3 and the domain D_2. The fuzzy membership function representing this polarity is strongly influenced by the ones representing the polarities of concepts C_1 and C_5 with respect to the domain D_2.

The propagation of the learned information through the graph is done iteratively where, in each iteration, the estimated polarity value of the concept x learned during the PL phase is updated based on the learned values of the adjoining concepts. At each iteration, the updated values are saved in order to exploit them for the re-shaping of the fuzzy membership function associating the concept x to the domain i.

The resulting shapes of the inferred fuzzy membership functions will be trapezoids where the extension of the upper base is proportional to the difference between the value learned during the PL phase (V_{pl}) and the value obtained at the end of the IP phase (V_{ip}), while the support is proportional to both the number of iterations needed by the concept x to converge to the V_{ip} and the variance with respect to the average of the values computed after each iteration of the IP phase.

3.3 Polarity Aggregation And Decision Phases

The fuzzy polarities of different concepts, resulting from the IP phase, are finally aggregated by a fuzzy averaging operator obtained by applying the extension principle (for the technical details see [11]) in order to compute fuzzy polarities for complex entities, like texts, which consist of a number of concepts and thus derive, so to speak, their polarity from them. When a crisp polarity value is needed, it may be computed from a fuzzy polarity by applying one of the defuzzification methods proposed in the literature [11].

Let $\mu_C : [-1, 1] \to [0, 1]$ be the fuzzy interval (i.e., a convex fuzzy set) representing the fuzzy polarity of concept C resulting from the IP phase. Let T be a text (or any other entity that may be regarded as a combination of concepts) related to concepts $C_1, \ldots, C_n$. The fuzzy polarity of T, $\mu_T : [-1, 1] \to [0, 1]$, may be defined as the average of the fuzzy polarities of concepts $C_1, \ldots, C_n$, by applying the extension principle, as follows, for all $x \in [-1, 1]$:

$$\mu_T(x) = \sup_{x=\frac{1}{n}\sum_{i=1}^{n} x_i} \min_{i=1,\ldots,n} \mu_{C_i}(x_i). \tag{2}$$

The result of the polarity aggregation phase is a fuzzy polarity, whose membership function reflects the uncertainty of the available estimate obtained by the system. In this sense, μ_T may be regarded as a possibility distribution of the actual polarity of T. Given $x \in [-1, 1]$, the membership degree $\mu_T(x)$ represent the degree to which it is possible that the polarity of T is x. Here, we are making the assumption that polarity is gradual, i.e., that a text may be more or less negative or positive.

At some point, if a decision must be made based on the polarity of T, some criterion has to be adopted, which takes the uncertainty of the estimate into account. As a matter of fact, a criterion can be defined only with reference to a given application scenario. For instance, if we can afford any desired number of texts and what we want is to pick a few of them whose polarity is certain, we can look for T such that either $d_T < 0$ or $a_T > 0$, i.e., the support of μ_T lies entirely on the left or on the right of zero, because in those cases it is certain that polarity is negative (in the former case) or positive (in the latter). In other scenarios, where what we want is to classify each and every text as either negative or positive as accurately as possible, we will have to be less picky and rely on a defuzzification method to transform μ_T into a crisp polarity value.

4 Challenge Tasks and Conclusion

In this paper, we have presented a fuzzy concept-based sentiment analysis system able to model fuzzy membership functions representing the polarities and the aboutness of concepts with respect to a particular domain. The system has been implemented in the context of the ESWC 2014 Concept-Level Sentiment Analysis Challenge. The Tasks proposed by the challenge have been addressed as follows.

Elementary Task: the polarity of each text is computed by aggregating the fuzzy membership functions associated with the extracted concepts. The aggregation operation is performed by applying the extension principle as described in Section 3.3.

Advanced Task #1 and #2: both aspects and concepts (simple and complex) are extracted by exploiting the built knowledge base (as explained in Section 2) and, concerning the Advanced Task #1, its polarity is computed by applying the approach used in the Elementary Task.

Advanced Task #3: similarly to the Elementary Task, the classification of each text is done by analyzing the associations between concepts and domains (independently from the polarity); therefore, the domain of each text is extracted by applying the extension principle of fuzzy sets.

Finally, the system has been preliminarily tested on the full version of the Blitzer dataset as shown in Table 1[7]. The system has been compared with three different baselines representing the most well-known machine learning techniques available today demonstrating the feasibility of the proposed approach for addressing the multi-domain sentiment analysis problem.

SVN [12]	Naive-Bayes [13]	Max-Entropy [13]	MDFSA Precision	MDFSA Recall
0.8068	0.8227	0.8275	**0.8617**	0.9987

Table 1: Results obtained on the full version of the Blitzer dataset.

References

1. Pang, B., Lee, L., Vaithyanathan, S.: Thumbs up? sentiment classification using machine learning techniques. In: Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), Philadelphia, Association for Computational Linguistics (July 2002) 79–86
2. Blitzer, J., Dredze, M., Pereira, F.: Biographies, bollywood, boom-boxes and blenders: Domain adaptation for sentiment classification. In Carroll, J.A., van den Bosch, A., Zaenen, A., eds.: ACL, The Association for Computational Linguistics (2007)
3. Bollegala, D., Weir, D.J., Carroll, J.A.: Cross-domain sentiment classification using a sentiment sensitive thesaurus. IEEE Trans. Knowl. Data Eng. **25**(8) (2013) 1719–1731
4. Xia, R., Zong, C., Hu, X., Cambria, E.: Feature ensemble plus sample selection: Domain adaptation for sentiment classification. IEEE Int. Systems **28**(3) (2013) 10–18
5. Ponomareva, N., Thelwall, M.: Semi-supervised vs. cross-domain graphs for sentiment analysis. In Angelova, G., Bontcheva, K., Mitkov, R., eds.: RANLP, RANLP 2011 Organising Committee / ACL (2013) 571–578
6. Tsai, A.C.R., Wu, C.E., Tsai, R.T.H., jen Hsu, J.Y.: Building a concept-level sentiment dictionary based on commonsense knowledge. IEEE Int. Systems **28**(2) (2013) 22–30
7. Fellbaum, C.: WordNet: An Electronic Lexical Database. MIT Press, Cambridge, MA (1998)
8. Kipfer, B.A.: Roget's 21st century thesaurus, third edition. (May 2005)
9. Cambria, E., Speer, R., Havasi, C., Hussain, A.: SenticNet: A publicly available semantic resource for opinion mining. In: AAAI Fall Symposium: Commonsense Knowledge. Volume FS-10-02 of AAAI Technical Report., AAAI (2010)
10. Liu, H., Singh, P.: ConceptNet—a practical commonsense reasoning tool-kit. BT Technology Journal **22**(4) (October 2004) 211–226
11. Zadeh, L.A.: The concept of a linguistic variable and its application to approximate reasoning - i. Inf. Sci. **8**(3) (1975) 199–249
12. Chang, C.C., Lin, C.J.: LIBSVM: A library for support vector machines. ACM TIST **2**(3) (2011) 27
13. McCallum, A.K.: Mallet: A machine learning for language toolkit. `http://mallet.cs.umass.edu` (2002)

[7] Detailed results and tool demo are available at `http://dkmtools.fbk.eu/moki/demo/mdfsa/mdfsa_demo.html`